APPENDIX

Descriptions of coalescent generators

We indicate in this section the general formulae for different types of coalescent models describing the rates of coalescence of k lineages when there are n active ancestral lineages in total. For a sample size n, each k-tuple of lineages is merging to form a single lineage at rate $\lambda_{n,k}$, and no other transitions are possible.

$$\lambda_{n,k} = \int_0^1 x^{k-2}(1-x)^{n-k}\,\Lambda(dx) \text{ with } 2 \leq k \leq n \qquad \text{[A1]}$$

Where Λ is a probability measure on [0,1] which defines the size, i.e. how many lineages merge, and probability, that is how often merging of lineages occur, of the coalescent events.

1) Kingman coalescent

The Kingman coalescent is obtained from eq. A1 with:

$\Lambda(\mathrm{dx}) = \delta_0\, dx$ [A2]

which is the Dirac point mass function at 0. In other words, the Kingman coalescent is defined with the rate of coalescence being $\lambda_{n,k}= 0$ except if k=2 where $\lambda_{n,2}$=1. Only binary collisions, *i.e.* coalescence of two lineages, are thus possible.

2) Λ-coalescent

The Λ-coalescent is generally defined in eq. A1 where Λ specifies the size and probability of the coalescent events following given specific Dirac mass functions with peculiar interpretations. The function

$\Lambda(\mathrm{dx}) = c\delta_y dx$ ($c > 0$) [A3]

defines a model where a fraction y of the population is replaced by the offsprings of a single ancestor at rate c. This macroscopic reproduction event is called an "extreme reproductive behaviour", and leads to multiple collisions in the genealogy (coalescence of more than 2 lineages).

Another special extreme case occurs when choosing $\Lambda = \delta_1 dx$ where the Dirac mass function is on one, which generates a star like-genealogy where all n lineages merge to a single ancestor.

3) Beta-coalescent

This class of coalescent is a special case of Λ-coalescents for which Λ has a density define as a function Beta(α,2- α) (with $1 \leq \alpha < 2$). Note that the special case where $\alpha = 2$ corresponds to the classic Kingman coalescent above.

$$\Lambda(dx) = \frac{1}{\Gamma(2-\alpha)\Gamma(\alpha)} x^{1-\alpha}(1-x)^{\alpha-1}dx \qquad [A4]$$

As for the Λ-coalescent, the “extreme reproductive behaviour” in eq. A4 leads to multiple collisions in the genealogy with coalescence of more than 2 lineages.

4) Bolthausen-Snitzman

The Bolthausen-Snitzman coalescent is obtained as a specific case of the Λ-coalescent where $\Lambda(dx)$ = Beta(1,1), or in other words, a beta-coalescent with $\alpha = 1$. In this case, the Dirac function has a uniform distribution of mass on [0,1].

5) ψ-coalescent

Eldon and Wakeley (2006) have developed a model where a single parent chosen at random from the population of size N contributes either 1) one offspring with probability 1-ε as in the classic Wright-Fisher model, or 2) ψN offsprings with probability ε. The latest case generates a large reproduction event, and a multiple merging of several lineages. The case where $\varepsilon = 1/N^{\gamma}$ is considered so that if $0 < \gamma < 2$. The function Λ from eq. A1 is:

$$\Lambda(dx) = (\delta_0 + c\psi^2\delta_\psi)dx$$

This function defines a model where a fraction ψ of the population is replaced by the offsprings of a single ancestor at rate c.

Eldon and Wakeley rewrote eq. A1 to define the rate of multiple merger of k lineages among n under their model:

$$\lambda_{n,k} = \binom{n}{k}\psi^k(1-\psi)^{n-k}$$, with $0 < \psi < 1$.